Journal of Imaging Science and Technology® 50(5): 469–480, 2006.

Methodology for Evaluating the Quality of ICC Profiles—Scanner, Monitor, and Printer

Abhay Sharma▲
School of Graphic Communications Management, Ryerson University, Toronto, M5B 2K3, Canada
E-mail: sharma@ryerson.ca

Abstract. *The concepts and technology for an open loop color management system such as that specified by the International Color Consortium (ICC) have been around for a number of years. The adoption of this workflow by the graphic arts industry has been slow. A major contribution to the lack of popularity is that the quality part of the ICC workflow is unregulated and the average user is unable to independently assess the quality of profiles and profile making software. This paper describes a number of test methods that can be used to evaluate the colorimetric accuracy of ICC scanner, monitor, and printer profiles. ICC profiles are being used in a number of color proofing scenarios. In order to understand the color reproduction abilities of such proofing systems it is necessary to quantify the accuracy of the underlying ICC profiles used in the workflow. A quality metric can be useful to provide feedback on how well a device has been characterized and therefore provide limits on the ability of a color managed system. Further, a universally defined merit figure will allow the comparison of results across manufacturers, allowing the user to make informed choices appropriate for their workflow. If we are able to establish a benchmarking procedure akin to the miles per gallon fuel consumption quoted for motor vehicles this provides a universal quality metric that can help raise the quality of profiling software, assist user choice, and ultimately lead to the greater acceptance of ICC color management in graphic arts and the printing industry. The quality of input profiles is described in terms of a* ΔE *calculation. Input profiles from ten different profiling packages were evaluated. Profiles were made for a flatbed scanner using Agfa, Fuji, and Kodak IT8.7/2 targets. The average* ΔE *for input profiles is shown to be in the range of* ΔE *0.60–2.46. A procedure is described for evaluating the quality of monitor profiles in terms of measured gamma, white point and the color reproduction of 24 specially chosen colors. Eleven monitor systems were evaluated. Monitor profiles were made for an Apple Cinema HD LCD display and measurements were made to verify the accuracy of these profiles using a telespectroradiometer. It is shown that commercial products were able to reproduce 24 patches of a color checker on a monitor with an average* ΔE *of 2.92–5.81. A printer profile metric is also described. This research describes three possible metrics for a printer profile—accuracy of the B2A1 tag (PCS to Device), A2B1 tag (Device to PCS), and a round trip test. Data are presented to show that the average accuracy of the output profile colorimetric intent (average of B2A1 and A2B1) for an Epson ink jet printer can be between* ΔE *1.72–3.49. The accuracy of the printer profile is useful when considering the use of an ink jet printer in color management proofing workflows. This research proposes a* ΔE *metric system that can be used to evaluate the quality of commercially created ICC input, display, and output profiles. The data presented here are fundamentally a methodology that can be used to estimate the colorimetric accuracy of profiles. The use of commercial profiling products is only to illustrate how the metrics may be determined in practice. The data, however, do show useful information about the state of color management products today and this analysis can be used to track the improvements and evolution in ICC profiling software.* © *2006 Society for Imaging Science and Technology.*
[DOI: 10.2352/J.ImagingSci.Technol.(2006)50:5(469)]

INTRODUCTION

The principles of an International Color Consortium (ICC) color managed workflow are now well established.[1–4] A color management system uses software, hardware, and set procedures to control color across different media. More specifically, an ICC color management system can be defined as a system that uses input and output profiles to convert device dependent image data into and out of a central, device independent profile connection space (PCS). Data in the PCS can be defined in terms of CIE *Lab* or CIE *XYZ*. Device characterization information is stored in profiles such that an input profile provides a mapping between input *RGB* data and the PCS, and an output profile provides a mapping between the PCS and output *RGB*/*CMYK* values.

Device characterization involves creating a mathematical representation of a device's color behavior. Characterization data are stored in the form of single and multidimensional lookup tables in a profile. The accuracy of color from input to the displayed image, to the printed image, depends in part on the quality of the characterization. A color management system will work well if the characterization data accurately represent the device's real color behavior. If there are issues with color reproduction in a workflow, a quality measure can be used to incriminate (or eliminate) the profile as a cause of the problem.

The growth in color management means that there are now many different software packages that can make ICC profiles. How do we know which is the most accurate and which represents the best value? It is important to establish a procedure for determining a meaningful merit figure for the accuracy of an ICC profile and for industry and researchers to agree on how this figure is calculated. It is suggested that following profile generation, profiling software should report a quality metric. This will provide information on how well a particular device has been characterized. In a turnkey ICC color-imaging situation involving unskilled personnel, a single merit figure can be used to devise a "go/no-go" decision making strategy. Some vendors quote a ΔE merit figure and often programs will write out a file with statistics. How-

▲IS&T Member

Received Oct. 17, 2005; accepted for publication Jan. 10, 2006.
1062-3701/2006/50(5)/469/12/$20.00.

ever, there is no indication to tell us how these figures are calculated and whether everybody is measuring the same thing in the same way.

The ICC's main document—the ICC Specification[5] describes in detail the format and data structure of an ICC profile. The specification does not, however, describe the contents of a profile in terms of quality or accuracy. The ICC Specification does not stipulate the contents of a profile, thus, individual vendors may populate the contents of a profile's lookup tables with any data they like as long as it technically conforms to the format of the specification. This leads to the situation where a profile can be labeled as "ICC compliant," however, there is no guarantee to the quality of the contents of the profile.

This research describes a series of test procedures that can be used to quantify the numerical accuracy of scanner, monitor and printer profiles. The aim of this research is to demonstrate a methodology for a base line numerical assessment for the colorimetric accuracy of ICC profiles, which can be used to assist user choice, raise the standard of profiling software, and thus promote ICC color management implementation.

Attempts have been made to evaluate the errors in ICC color reproduction and results are described in the literature for the analysis of end-to-end errors in a soft-proofing system[6] and the comparison between the accuracy of ICC profiles in relation to proprietary style files.[7] Results have also been published to quantify the errors in digital camera characterization.[8,9] The accuracy of the color conversion depends on many issues that may be considered external to the profile itself. These include the choice of PCS (CIE *XYZ* or CIE *Lab*), the choice of color management module (CMM) and the way that the CMM may concatenate the profile one-dimensional and multidimensional lookup tables. Many of the issues that affect the quality of processed images using ICC color management protocols have been described.[10,11] Adams and Weisberg[12] conducted an experiment similar to that described in this work on nine profile making software packages and other studies have measured the accuracy of the input profile.[13] Berns has described the accuracy of the color reproduction when using ICC profiles to reproduce images in a color text book.[14] Nevertheless, there still exists the need to establish a procedure and a universal, simple, and meaningful merit figure for measuring the accuracy and quality of ICC profiles.

The assessment of ICC profiles and color reproduction is a complex issue involving everything from color science, psychophysics, and image analysis to "preferred" reproduction styles. The approach adopted in this work is to evaluate the accuracy of a profile using the colorimetric intent. This does not provide an all-encompassing result but does provide an indicative set of base line metric figures that can be used to make valid cross-vendor comparisons.

While colorimetric accuracy is important in profile evaluation, the perceptual intent, smoothness, monotonicity, gray scale reproduction, and other variables should also be considered. In addition to the tests described here for colorimetric accuracy it would also be necessary to test other aspects of a profile. Tests for the perceptual part of a profile can be conducted using MATLAB routines[15] and tests for smoothness have been described.[16,17] Researchers have also studied color discontinuities due to color contouring as a two-dimensional problem.[18]

This research describes a metric system and uses it to evaluate the quality of a number of commercially available profiling packages. The procedures are applied to ICC profiles for a flatbed scanner, LCD display, and an ink jet printer. First, we look at the results for an input profile. We describe why an input profile may have errors (fit issues) and then suggest a mechanism to calculate a ΔE metric. Where possible we compare the manufacturers quoted metrics to those computed in this study. A metric for monitor profiles is then described. Where appropriate, the same set of vendors that were used for input profile analysis is used for assessing monitor-profiling accuracy. Finally, printer profiles are evaluated. For each product tested default settings were used, no attempt was made to alter the vendor's starting recommendations.

Metrics used in the paper for reporting color accuracy are mean and maximum ΔE. These are overall metrics convenient to compare a long list of experimental cases, however, in many instances, for example when assessing colors, or trying to identify the source of errors, it may be necessary to calculate separately the color difference as ΔL^*, Δa^*, or Δb^*. It is also useful to know if the difference between the samples comes mainly from a difference in lightness (L^*) or chroma (a^*, b^*), as this is informative about color casts in the reproduction.

In some instances, first-order statistics of mean and maximum ΔE for the comparison of two colors is not adequate. Second-order statistics using histograms and cumulative frequency distributions provide a better indication of the overall color match.[19,20]

It should be noted that while ΔE is a useful measure it is not an all inclusive quality measure and it would generally be necessary to consider image content issues as well. A "busy" image may mask color errors while a single large color patch may accentuate errors. Due to the imperfections in the CIE *Lab* system, errors in midneutral gray may be more objectionable and easier to notice than the same error in a dark saturated color. Traditional $\Delta E(a^*b^*)$ was used in all cases, however, to improve the correlation between human perception and numerical figures it may be beneficial to consider newer ΔE metrics. Image dependent analysis of profile quality is implicitly contained within the suggested quality metrics but not explicitly examined.

It should be noted that color management software is continuously changing and improving and the commercial data presented here is merely presented to show some indicative numbers and a benchmark or snapshot of the industry at the present time. It is inevitable in a survey of this type that some vendors fared better than others, however, this should not be taken as an endorsement of any particular product or manufacturer; these numbers are only provided to demonstrate the process for today's commercially available products. Testing was conducted in January 2005.

The data in this report are subject to instrumental vagaries including device repeatability, interinstrument agreement and accuracy of the measuring device.[21] Device repeatability introduces an "error bar" to the data. Interinstrument agreement is avoided as a single instrument is used in all the tests. Instrument accuracy is explicitly tested in the case of monitor profiling. In the section on monitor profiling, different vendor measurement devices are used to measure the display and construct a profile. A separate, high-quality measuring instrument is then used to evaluate the quality of the profile and thus implicitly the accuracy of the measuring instrument.

In this research the following types of devices were used—Umax Astra 4000u (low-end scanner), Apple 23 in. Cinema HD LCD Display, and Epson Stylus Pro 4000 (CMYK inkjet printer with Ultrachrome inks and semimatt proofing paper).

First the test procedure for scanner, monitor and printer profiles is defined, then it applied to the following commercial products—ColorSolutions basICColor™, Digital Light & Color Profile Mechanic™, Fujifilm ColourKit Profiler Suite™, GretagMacbeth Eye-One Match™, GretagMacbeth ProfileMaker™, Heidelberg PrintOpen™, Pantone ColorVision Spyder2PRO™, QPI ColorBlind Pro™, TGLC PerfX Color Management™, X-Rite MonacoEZColor™, X-Rite MonacoOPTIX™, and X-Rite MonacoProfiler™. Generic profiles and Apple's Display Calibrator utility were also tested. ICS RemoteDirector is not intended for stand-alone profiling and GMG ColorProof is a proprietary, non-ICC proofing program. Also a printer profile was made using a remote profiling service called ColorValet™ by Chromix. All testing was done with Mac OS 10.3.6, Photoshop CS and the ACE CMM, except for Heidelberg PrintOpen and ColorBlind that are Windows programs. Most users have placed little importance on version 4 ICC profiles, and while most vendors will make version 4 profiles this is not the default, so in this work version 4 profiles were not used.

INPUT PROFILE ACCURACY

An input profile in a color-managed system provides a transformation between scanner *RGB* and device independent CIE *XYZ* or CIE *Lab*. The process of generating and storing this transform is called characterization. To construct the transform a scan is made of a standard characterization test chart to obtain scanner *RGB* values. The test chart patches are also measured using an instrument such as a spectrophotometer to provide corresponding *Lab*/*XYZ* colorimetry. A mathematical relationship is then derived between the scanner *RGB* values and corresponding *Lab* data. This transform information is stored in (ICC standardized) single and multidimensional lookup tables. These lookup tables constitute the main component of an ICC input profile.

What does the accuracy of the input profile depend on and why can some vendors get better results than others? A major part of the accuracy of the profile depends on the model used to determine the relationship between scanner *RGB* and corresponding *Lab* or *XYZ* values. The literature describes a number of different ways to establish this transform relationship. It is possible to use data fitting processes that can range from a simple linear matrix approximation to higher order polynomial regression.[22] Due to the nonlinear relationship between dye density and tristimulus value, charge coupled device flatbed scanners that are primarily designed to measure densities are poorly characterized by a linear transformation.[23] The transform between scanner *RGB* and *Lab* is therefore most commonly computed using polynomial regression. It may be necessary to use a higher-order polynomial least squares fit process to adequately characterize the scanner response. A least squares fit process solves simultaneous equations and determines a set of polynomial coefficients that relate scanner *RGB* to *Lab*. Because the polynomial fit process attempts to satisfy all the training set data, it is subject to fit errors even for the training points. A higher-order polynomial can introduce erratic results outside the training set region where interpolation or extrapolation occur, producing local maxima and minima which lead to problems in processed image colors. The order of the polynomial needs to be carefully chosen so as to maximize colorimetric accuracy without introducing unwanted artifacts. To create a better fit to the data, polynomial regression analysis can be used in conjunction with some prelinearization[24] and often it is found that mapping *RGB* to *XYZ* is preferable to mapping *RGB* to *Lab*.

From the above it is obvious that constructing an input profile characterization involves more than simple mathematical fitting of two data sets. It is necessary to provide an accurate colorimetric transform not only for training set data but also for every other possible *RGB* triplet combination. Different manufacturers deal with the constraints on the process in different ways, this is part of the reason that the measured accuracy may vary between vendors.

In the tests conducted as part of this research, scanner profiles were made with manufacturer supplied batch-measured *Lab* reference data files. Another method to obtain the *Lab* reference file is for a user to custom measure the target. It should be noted that the accuracy of the process is not dependent on using batch-measured or custom-measured targets.[25] To appreciate this last statement consider the process under discussion. An input profile provides a transform between *RGB* and *Lab*. Custom and batch measured *Lab* data may be slightly different which can be visualized as two slightly shifted clouds of *Lab* values. The characterization process seeks to create a transform between some fixed *RGB* values and one or other of the *Lab* clouds. As long as the target has not faded or deteriorated in an "irregular" manner the *Lab* clouds are just displaced and the ability of the profile making software to create an accurate transform is largely unchanged by the chosen *Lab* data set. Note that there will, however, be a difference in the accuracy of scanned imagery depending on whether batch or custom data are used to make the profile. During scanning of images the overall reproduction of the original colors to reproduced colors will be more accurate if the profile is made from custom measured data.

INPUT PROFILE—TEST PROCESS

A Δ*E* metric for input profile accuracy is proposed. A general description of the metric evaluation process is that a scan of a standard test chart is used to make an input profile and then the same image is used again to test the profile and derive a Δ*E* metric. The first part of the process is to make an input profile in the normal way. To construct an input profile a scan is made of the standard characterization test chart to obtain scanner *RGB* values. The reference file for the test chart containing corresponding *XYZ*/*Lab* values is obtained. The scan of the chart and the reference file are provided to a commercial profile making package that computes the mapping transform between *RGB* and *Lab*, populates the lookup tables and saves the result as an ICC input profile.

To compute a Δ*E* accuracy metric, the *RGB* values of the scanned chart image are processed through the input profile to arrive at processed *Lab* values. A program such as Adobe Photoshop™ can be used to do this. The processed *Lab* values are compared to reference *Lab* values. Ideally, the processed data should be numerically equivalent to the reference data. Due to fitting processes and interpolation errors, there is likely to be a difference between these two values. A Δ*E* difference can be calculated between the processed *Lab* data and the reference *Lab* data and this forms a metric for input profile quality. The Δ*E* number reflects the goodness of the vendor's underlying model for the relationship between *RGB* and *Lab*, quantization errors in the lookup table and any CMM concatenation errors. The Δ*E* metric provides a visually relevant measure of the magnitude of color difference and is indicative of the likely errors that will be encountered in the workflow when the profile is applied to (or associated with) scanned imagery. This simple result is a useful guide to the accuracy of the input profile and is a useful metric that can be used to assess the relative quality of input profiles from different sources for the same data set.

INPUT PROFILE—RESULTS

To demonstrate the test process, the Δ*E* error for a scanner profile was calculated for a number of commercial profiling programs. An IT8.7/2 reflection test target was scanned on a Umax Astra 4000u scanner with all image correction controls turned off. The corresponding *Lab* batch-measured reference file was located and an input profile was made using a number of commercially available profile making software packages as listed in Table I. Each program was presented with the same target scan and reference file. Additionally, a generic profile was obtained as part of the Umax scanner driver, Umax VistaScan 3.5.4. In making the scanner profiles, default options within the profiling software were chosen. The input profile was made and saved and any error figures provided by the manufacturer were noted. The process was repeated for different IT8.7/2 targets so that at the end of the experiment, profiles were made using each manufacturer's software with an Agfa (Agfacolor paper, 1999:03), FujiFilm (Fujicolor paper, 2000:05), and Kodak (Ektacolor paper, 1997:04) reflection IT8.7/2 targets.

Table I. Results for the quality of the scanner profile with three different chart types.

	Agfa IT8.7/2 chart	Fujifilm IT8.7/2 chart	Kodak IT8.7/2 chart	Final result
Scanner profile quality Umax Astra 4000u	Mean (Max) Δ*E*	Mean (Max) Δ*E*	Mean (Max) Δ*E*	Average Δ*E*
X-Rite Monaco Profiler 4.7	0.67(9.66)	0.50(3.87)	0.63(6.17)	0.60
X-Rite Monaco EZColor 2.6.3	0.70(8.63)	0.53(4.51)	0.63(6.11)	0.62
Fujifilm ColourKit Profiler Suite 4.2	0.99(5.06)	0.87(3.85)	0.83(4.62)	0.90
TGLC PerfX Color Management 1.2.8	0.95(4.11)	1.06(4.01)	1.01(4.77)	1.01
GretagMacbeth Eye-One Match 3.0	1.09(3.94)	0.90(5.80)	1.19(5.33)	1.06
GretagMacbeth ProfileMaker 5.0.1	1.08(3.91)	1.15(15.14)	1.20(4.94)	1.14
Digital Light & Color Profile Mechanic 1.0.0.3	1.09(7.19)	1.00(5.06)	1.37(6.19)	1.15
QPI ColorBlind Pro 5.1 (Windows)	1.60(6.94)	1.90(9.49)	1.37(8.16)	1.62
ColorSolutions basICColor scan+2.2	2.37(9.29)	2.56(7.99)	2.46(10.78)	2.46
Generic Umax Scanner Profile	29.76(44.32)	28.85(42.01)	29.33(46.72)	29.31

To measure the accuracy of the scanner profile the following test was conducted. Following profile generation, the raw *RGB* scan of each IT8.7/2 chart image was opened in Adobe Photoshop CS. Each scanner profile was selected in turn using Image>Mode>Assign Profile and the image was processed to *Lab* using Image>Mode>Convert to Profile where the Destination Space was chosen as *Lab* Color. The rendering intent chosen was Absolute Colorimetric and the CMM used was Adobe (ACE). Next, a special program, written in our laboratory was used to average the central portion of each patch. Averaging was conducted over approximately 30×30 pixels from the center of each patch. The *Lab* value of each patch in the chart image was recorded in a text file. GretagMacbeth MeasureTool was then used to compute the Δ*E* between this value and the original reference value used in profile creation. A mean and maximum Δ*E* was noted for all patches of the IT8.7/2 target. For each vendor the test was repeated using Agfa, Fujifilm, and Kodak targets, and an average of the mean Δ*E* for the different charts was calculated and is shown in Table I.

Table II. Evolutionary data for the quality of the scanner profile for commercial products.

	Agfa IT8.7/2 chart	Fujifilm IT8.7/2 chart	Kodak IT8.7/2 chart	Final result
Scanner profile evolution	Mean (Max) ΔE	Mean (Max) ΔE	Mean (Max) ΔE	Average ΔE
Fujifilm ColourKit 2.2	1.17(3.98)	1.25(4.53)	1.42(3.66)	1.28
Fujifilm ColourKit 2.3	1.15(3.72)	1.23(4.53)	1.43(3.53)	1.27
Fujifilm ColourKit 3.0	1.11(4.36)	0.90(3.52)	0.88(4.47)	0.96
Fujifilm ColourKit 4.2	0.99(55.06)	0.87(3.85)	0.83(4.62)	0.90
Gretag ProfileMaker 3.1	0.85(2.59)	0.97(3.21)	1.16(3.30)	0.99
Gretag ProfileMaker 4.0	0.85(2.87)	0.99(10.13)	1.23(4.12)	1.02
Gretag ProfileMaker 4.1	1.15(3.59)	1.12(2.86)	1.22(4.91)	1.16
Gretag ProfileMaker 5.0	1.08(3.91)	1.15(15.14)	1.20(4.94)	1.14
Monaco Profiler 3.2	4.39(15.00)	5.04(8.25)	4.79(11.35)	4.74
Monaco Profiler 4.0	1.19(9.95)	0.92(4.70)	1.19(7.10)	1.10
Monaco Profiler 4.5	1.25(11.31)	0.91(4.40)	1.19(9.02)	1.12
Monaco Profiler 4.7	0.67(9.66)	0.50(3.87)	0.63(6.17)	0.60

As a check, the above process was also duplicated using an Apple supplied Applescript routine called "Match to chosen profiles." It was confirmed that the Photoshop process described above and the Applescript routine produced essentially identical results.

The accuracy of each vendor's product is shown in Table I. A lower ΔE number is preferable. Based on this table we see that many vendors create very accurate scanner profiles, shown by an average $\Delta E < 2$. In each case the maximum ΔE should also be considered. While it is desirable to have a low mean, it is also necessary to have no high maximum ΔE. A high maximum ΔE indicates that the profile will not accurately reproduce some specific color groups. The best result would ideally be a low mean and a low maximum ΔE. It is also important that each profiling program should be able to make an accurate profile with the Agfa, Fujifilm, or Kodak targets. Each emulsion type is different and the underlying mathematical model employed by the manufacturer should be robust so as to provide an accurate transform between *RGB-Lab* pairs from different film types. In a few cases the results were different across the emulsion types. Fujifilm ColourKit Profiler Suite 4.2, TGLC PerfX Color Management 1.2.8, and GretagMacbeth Eye-One Match 3.0 performed well as they all had a low mean and a low maximum ΔE across all chart types.

Users often ask—how good is the generic profile supplied by the manufacturer? For this scanner the generic profile with a ΔE of nearly 30 was very poor. Note that just because the generic profile is poor, this does not mean that the Umax scanner is poor. In fact the scanner is remarkably good value, the ΔE metric merely tells us that the generic profile, in this case, does not provide an accurate representation of the scanner's color characteristics.

Table I can be used for other purposes. Analysis of the results suggest that some vendors may be using the same core for consumer and professional versions of their software. For example note that MonacoEZColor and MonacoProfiler produce similar results, also GretagMacbeth's Eye-One Match and ProfileMaker are also next to each other in the ranking. We could conclude that these companies are using the same code in both their products.

ICC profiles can contain different look-up tables for different rendering intents—A2B0 (perceptual), A2B1 (colorimetric) and A2B2 (saturation). However, this was not always the case. In the early ICC File format specification, scanner and monitor profiles used to have only one lookup table, which was called the A2B0 tag. In the 1998 ICC specification, the A2B1 and A2B2 tags for the scanner profile were mentioned but were "undefined." Since the version 4 revision of the ICC specification the A2B0, A2B1, and A2B2 tags for all profiles are explicitly defined.[5] All profiles can now have the A2B0, A2B1, and A2B2 tags, thus there is no excuse for vendors to place colorimetric data (A2B1) in the perceptual (A2B0) tag or vice versa. It is interesting to note that the default behavior of GretagMacbeth ProfileMaker 5 is to make a scanner profile in which the colorimetric lookup table tag (A2B1) contains the contents of the perceptual lookup table (A2B0). To avoid any confusion it is recommended that vendors populate lookup tables in complete accordance with the ICC specification and that Adobe Photoshop is unambiguous in its use of rendering intents in all parts of the workflow.

INPUT PROFILE—HISTORICAL ANALYSIS

Using the current results together with previously published data[26–28] it is possible to conduct some historical analysis. It is possible to track the accuracy of profiling software during version change evolution. From the data in Table II, we could conclude that Fujifilm ColourKit was not changed between versions 2.2 and 2.3 but has been improved in version 3.0 and 4.2. GretagMacbeth's ProfileMaker has been variable over its version history. We may conclude that the code for MonacoProfiler was greatly improved between versions 3.2 and 4.0, versions 4.0 and 4.5 were essentially the same, and version 4.7 shows improvements that reduce the ΔE error by 50%. Note that there may be improvements in these products that are not detected by this test and that in scanner profiling it is possible to get a slightly different result each time the experiment is conducted due to user defined cropping of the test chart.

INPUT PROFILE—VENDOR PREDICTION

In many color management products the user is provided some feedback following profile generation. Table III shows how the vendor's prediction compares with the results calculated in this research. The close correlation between the data calculated here and the vendor's predictions suggests

Table III. Results for the quality of the scanner profile compared to the vendor's prediction.

Scanner profile quality Umax Astra 4000u		Agfa IT8.7/2 chart Mean (Max) ΔE	Fujifilm IT8.7/2 chart Mean (Max) ΔE	Kodak IT8.7/2 chart Mean (Max) ΔE	Final result Average ΔE
X-Rite Monaco Profiler 4.7	Calculated	0.67 (9.66)	0.50 (3.87)	0.63 (6.17)	0.60
	vendor's prediction	0.57 (9.20)	0.47 (4.00)	0.52 (6.10)	0.52
Fujifilm ColourKit Profiler Suite 4.2	Calculated	0.99 (5.06)	0.87 (3.85)	0.83 (4.62)	0.90
	vendor's prediction	0.87 (5.44)	0.77 (3.67)	0.79 (3.49)	0.81
TGLC PerfX Color Management 1.2.8	Calculated	0.95 (4.11)	1.06 (4.01)	1.01 (4.77)	1.01
	vendor's prediction	1.15 (3.55)	1.18 (4.13)	1.15 (5.12)	1.16
Digital Light & Color Profile Mechanic 1.0	Calculated	1.09 (7.19)	1.00 (5.06)	1.37 (6.19)	1.15
	vendor's prediction	1.11 (8.09)	1.06 (4.94)	1.45 (6.28)	1.21

that vendors are doing a calculation essentially similar to that described in this work. This is highly desirable. Vendors have the option of using their internal (prelookup table) mathematical model for computing the accuracy metric. The ΔE prediction using the prelookup table analytical model usually produces a better accuracy estimate, however it is unrealistic as it is not what the user will experience in practice. End-users do not have access to the underlying mathematical model and are forced to use quantized ICC input profile lookup tables that are subject to interpolation errors. If vendor's predictions agree approximately with the data produced in this research then we can assume that vendors are using a realistic (lookup table based) error calculation procedure.

INPUT PROFILE—TRAINING DATA VERSUS TEST DATA

In the input profile test procedure described above the scanned image data were used for both creating the profile and testing the profile. Thus the same data set was used for both training and testing.[29] It is often suggested that a profile will exhibit better accuracy for training set data, and data points that are different (slightly adjacent *RGB* values) to the training points will produce larger errors. Thus it is argued that using the same data for training and testing can create an estimate of accuracy that is inflated, i.e., that is better than reality. This argument sounds logical, but in practice it is not true. Due to the use of regularly spaced lookup tables in an ICC profile, using the same data set for training and testing does not artificially increase the accuracy of the profile. In the following discussion we explain why this is the case and present data to support the argument.

Figure 1 shows a graphical representation of the A2B lookup table of a scanner profile. In the figure, *RGB*(*s*) refers to the *RGB* values that are obtained when a test chart is actually "scanned." *RGB*(*l*) refers to the *RGB* values that form the nodes of the "lookup table." In the figure, the lookup table *RGB* node value is shown within the grid and the *Lab* "content" of the node is shown above the grid. The *Lab* value is obtained from the appropriate chart reference file. For clarity, only some *Lab* values are shown. All values are fictitious.

The A2B table can be used to lookup the *Lab* value for any given *RGB* value and *RGB* and *Lab* values that are in between nodes are determined by interpolation. It will be noted that due to the practical processes involved the *RGB*(*s*) values will not be regularly spaced. So if a lookup table is formed using the raw scan data, the cube nodes will be irregularly spaced. In theory the lookup table between irregularly spaced *RGB*(*s*) and *Lab* is an accurate representation of the scanner response and as such is a perfectly valid mathematical representation of the situation. However, in most color image processing applications including ICC color management systems, data are evenly distributed along each axis. What is required is reformatting of the information to present it on a regular grid. Most color management products will create an internal mathematical model representative of the device, and from this analytical description of the process data are extracted to populate the regularly spaced cube nodes. Color management software will take in *RGB*(*s*) and *Lab* data and create an input profile with an A2B lookup table based on *RGB*(*l*) and some new *Lab* data. Due to this process there exists little direct influence of a particular data point in *RGB*(*s*) data. Even if a target was designed in such a manner, despite adopting best practices, the *RGB* values in the scanned image are extremely unlikely

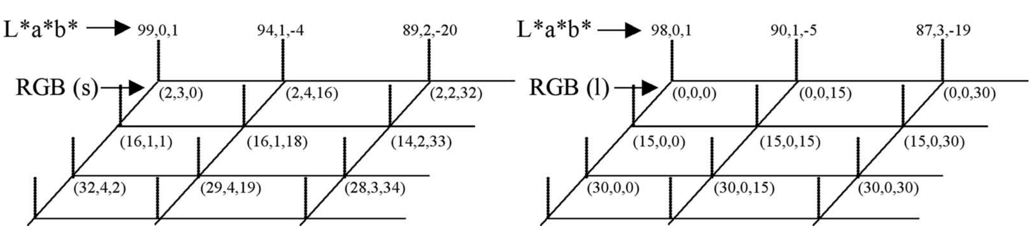

Figure 1. The original scan of the target produces irregular *RGB*(*s*) values that are processed to create a new regularly spaced lookup table, *RGB*(*l*).

Table IV. Quality of a scanner profile evaluated using testing versus training data.

	Profile tested with	
Mean (Max) ΔE	Kodak IT8.7/2 chart (1998-07)	Kodak IT8.7/2 chart (1997-08)
Scanner profile made with Kodak IT8.7/2 chart (1998-07)	1.76 (5.11)	1.87 (5.52)

to be the same as the *RGB* node values that form the *RGB*(*l*) to *Lab* lookup table.[7]

To test the above hypothesis a profile was made with one chart, in this case a Kodak 1998 chart. It was tested with its own data as a control, and then another data set of similar material was used. The results shown in Table IV suggest that there is not a significant difference between the accuracy of a profile when evaluated using testing versus training data.

MONITOR PROFILE ACCURACY

In scanner profiling there is a single, straightforward ΔE number that can be used to indicate the colorimetric accuracy of the profile, in monitor profiling, however, there is no unique profiling target and therefore there is no obvious test that is universal to all vendors. This research examines some of the parameters that need to be considered in order to evaluate the accuracy of a monitor profiling system.

Building a monitor profile involves use of a measuring device combined with a software package. All monitor profiling packages come with a bundled instrument. While it is possible to use different instruments with different software, in this testing the manufacturer's software-hardware combination was used. In monitor profiling there are good reasons to test the hardware-software as a pair. Monitor profiling packages typically use a colorimeter to reduce the cost of the product and thus appeal to wider market, such as digital photographers. The ability of the profile to accurately represent the color characteristics of the monitor is thus dependent on the accuracy with which the monitor is colorimetrically measured coupled with the quality of the profile generation. Thus in this testing the manufacturer's software-hardware combination is used and the profile quality metrics thus encompass both the accuracy of the measuring device and the ability of the vendor to create an accurate profile in software. In real life, users will use the hardware-software package, thus the tests described here are relevant to practical situations.

Monitor technology is simpler to characterize than scanners or printers. The response of a monitor is generally characterized by a linear expression (the phosphor matrix) combined with a nonlinear expression (the gamma curve).[30] Tags within a monitor profile represent both these parameters. With monitor profiles, a distinction can be made between characterization and calibration. In this context, characterization refers to the process where the monitor profile is used to simply represent the current state and behavior of

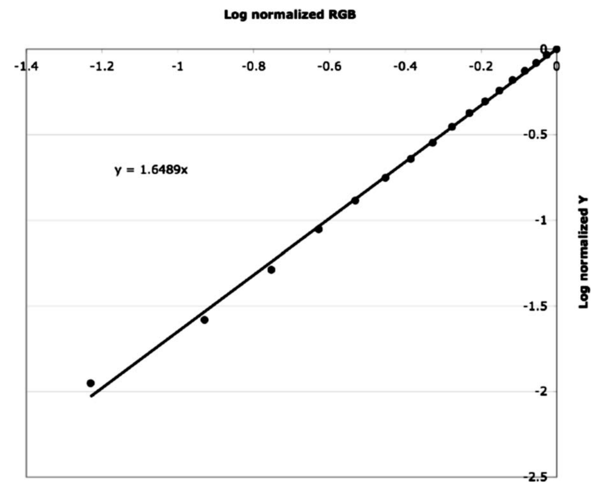

Figure 2. Monitor gamma can be easily determined as the slope a straight line in a log-log graph.

the device. The monitor profiling process can often be extended to first calibrate and then characterize. By this we mean that the profiling software can be used to calibrate (adjust the response of the monitor to some predetermined condition, i.e., a chosen white point and gamma) and then information regarding this new condition is saved in the monitor profile in a process we call characterization. All Macintosh systems are capable of calibration and characterization, not all Windows systems are capable of calibration and characterization.

Macintosh monitor profiles are distinguished by the use of the "vcgt" tag that is used to provide the calibration part of the system. Vcgt stands for video card gamma tag and has been part of Mac OS since ColorSync 2.5.[31] How is calibration to a user defined gamma and white point achieved? First the inherent, factory response of the system is determined. Then the software asks the user for the required gamma and white point. A correction is calculated and stored in the vcgt, such that the vcgt in conjunction with the factory response results in the user requested gamma and white point.[4] On selection of a monitor profile, the data from the vcgt tag is downloaded to the video card and used to actively alter the system display.

MONITOR PROFILE—TEST PROCEDURE

A number of profiling packages were tested to see if they were able to achieve a requested gamma, a requested white point and accurately reproduce 24 colors that approximately represent a Macbeth ColorChecker chart. Monitor profiles were made using different measuring instruments as shown in Table V. Where offered, the user requested a gamma of 1.8 and white point of D_{50}.

After each profile was made it was selected as the system profile. Using Photoshop CS, a grayscale ramp was displayed on the monitor consisting of *RGB* (0, 0, 0), (15, 15, 15), …, (255, 255, 255). The luminance (*Y*) was measured using a PR-650 SpectraScan spectrocolorimeter and a log-log plot was used to determine the gamma of the display by fitting a straight line to the data and noting the slope as the gamma value, Fig. 2.

Table V. Results for the quality of the monitor profile.

Apple 23″ Cinema HD Monitor profile quality	Measuring instrument	Achieved gamma (target was 1.8)	ΔE difference in white point from a target of D_{50}	Average ΔE of 24 patch Macbeth ColorChecker
Digital Light & Color Profile Mechanic Monitor 1.0	Sequel G4 CL	1.80	5.83	2.92
GretagMacbeth ProfileMaker 5.0.1	Spectrolino	1.75	1.99	3.05
ColorSolutions basICColor display 3.03	Eye-One	1.72	3.37	3.20
Integrated Color Solutions Remote Director 2.6.3	Eye-One	1.97	4.26	3.35
Fujifilm ColourKit Profiler Suite 4.2	Eye-One	1.70	4.31	3.57
GretagMacbeth Eye-One Match 3	Eye-One Display 2	1.70	5.64	4.15
Monaco Profiler 4.7	MonacoOPTIX^{XR}	1.71	6.14	4.22
Pantone ColorVision Spyder2PRO 1.0	Spyder2PRO	1.76	7.12	5.14
Apple Display Calibrator Assistant 4.2	None	1.65	6.86	5.55
MonacoOPTIX Pro 2.03	MonacoOPTIX^{XR}	1.70	9.71	5.56
Monaco EZColor 2.6.3	MonacoOPTIX^{XR}	1.71	9.61	5.81

Next a white patch of *RGB* 255, 255, 255 was displayed and the *XYZ* values of this patch were measured. The measured *XYZ* values were normalized to $Y=100$ (the color temperature is unchanged by a uniform rescaling of the *XYZ* values). The measured *XYZ* was converted to *Lab* for the chosen illuminant, D_{50} and compared to an ideal D_{50} white point that has an *Lab* of 100,0,0. A $\Delta E_{\mathrm{a,B}}$ calculation was done to establish how close each profile was able to create the requested color temperature. A $\Delta E_{\mathrm{a,b}}$ figure was defined as:

$$\Delta E_{a,b} = \sqrt{a^2 + b^2} = C.$$

Thus, we see that the $\Delta E_{\mathrm{a,b}}$ has a simple interpretation as the chroma, *C*, of the measured white point, referenced to the target white point. Finally, data for $Lab(D_{50})$ values representative of a 24 patch Macbeth ColorChecker were displayed in Photoshop. A PR-650 SpectraScan colorimeter was used to measure the *XYZ* of the patches, which was converted to $Lab(D_{50})$ and the ΔE was calculated and averaged over 24 patches. There was one patch that was out of gamut of the display and this produced an expected high error reading for many of the products evaluated.

MONITOR PROFILE—RESULTS

The results, Table V, show that many commercial profiling systems may be moving away from the traditional gamma for the display, which has arisen historically from the characteristics of a CRT display. Early configurations for LCD displays attempted to mimic the gamma response characteristics of a CRT on a LCD in order to make the two technologies more similar. The data presented here suggest that some vendors may be aiming for a gamma that is linear in L^* and a log-log fit for the gamma value may not be the best way of estimating the gamma characteristics of a LCD display. A newer (dynamic) approach to monitor characterization has been described.[32]

For the white point, a lower ΔE is better. For the colors of the Macbeth ColorChecker a lower ΔE is better and a ΔE of around 3 or less is likely to produce good results. Keep in mind that from scrutiny of the data it was obvious that at least one of the chosen colors was out of the gamut of the display. In the table pay attention to the instrument used, some instruments are colorimeters (Sequel G4 CL, Eye-One Display 2, MonacoOPTIX, and Spyder2PRO) and some are spectrophotometers (Spectrolino, Eye-One). Note that profiles were made with the different instruments listed in Table V but measurement was done with a totally different but single instrument (PR-650 SpectraScan). Remote Director 2.6.3 is not (strictly speaking) a program intended for monitor profiling but could be expected to achieve better results than ICC products as it is in total control of the display system—profile, CMM, video card, etc. RemoteDirector occupies a side-by-side ranking with ColorSolutions basICColor display—with which it shares a common ancestry. We should point out that the Apple Display Calibrator is a simple utility that makes a valid monitor profile but is not generally used in commercial workflows due to the fact that it is based on the user's visual assessment. Procedures for the "calibration by eye" method are described in the literature.[33]

The tests described here for monitor profile accuracy provide a necessary first level series of tests. In practical situations it would be useful to extend the tests to areas such as smoothness and clipping. A profile with a very low ΔE but with no shadow detail or with posterization in smooth gradations is likely to be less appealing to many users than a profile with higher ΔE but with no clipping and no posterization. The lookup table that is loaded into the video card is responsible for achieving the desired white point, achieving the desired gamma, maintaining neutrality, maintaining smoothness, and prevention of clipping. Some of these goals can work against each other. For example achieving the desired gamma and preventing clipping presents a tradeoff. A good measure of smoothness is hard to come up with since a non-smooth table can simply be correcting a problem with the physical response of the monitor.

PRINTER PROFILE ACCURACY

An important aspect of this work is to evaluate the quality of each profile separately. In a typical workflow, for example, a source (scanner) and destination (printer) profiles may be simultaneously applied to an image, in this instance errors in the overall reproduction are contributed by both profiles. This research evaluates each profile separately, so if we have color problems we are able to implicate the scanner or printer profile. In this section we look at the accuracy of the output profile. In this section we continue the philosophy of "modular testing" that identifies the accuracy of each individual part of the process. In a printer profile we make a distinction in the specific parts of the profile being tested.

ICC profiles contain tags for different rendering intents—perceptual, colorimetric, and saturation. Each intent has two parts—a B2A (Profile Connection Space to Device) and a A2B (Device to Profile Connection Space) lookup table. Tests were done to evaluate the accuracy of the lookup table containing data pertaining to the absolute colorimetric intent, and we measure separately the accuracy of the A2B1 and B2A1 tags of the output profile.

While some image workflows may use the perceptual intent, the colorimetric intent is used during the facsimile reproduction of images, during soft proofing when images are evaluated on a monitor and during proofing when press images are "returned" to the Profile Connection Space and printed on a proofing device. The colorimetric intent may also be used when legacy *CMYK* images are repurposed. So although the colorimetric intent is not always used to process photographic images it is used in many significant ICC workflows and as such is an easily calculated profile accuracy measurement.

To evaluate the B2A1 part of an output profile some known *Lab* values are converted to *CMYK* using the profile and CMM to be evaluated. The *CMYK* values are printed and the *Lab* of each patch is measured. The measured *Lab* is compared to the *Lab* that is being sent to the printer. The ΔE between the known *Lab* and the measured *Lab* is calculated and averaged. In such a test it is possible to have ΔE errors due to two sources. Some colors may be out of gamut of the destination device and the colors will necessarily be altered by gamut clipping, producing a large calculated ΔE. While we do not to include out of gamut colors in the colorimetric analysis, it would normally be very relevant to evaluate how a vendor has mapped out of gamut colors. Methods based on psychophysical experimentation would be most appropriate to evaluate image quality for out of gamut color mapping. The other source of errors is due to the inaccuracies of the profile and concatenation process. The out of gamut issue is due to the fundamental limitations of the device and to better concentrate on the error in the profile, we seek to remove the gamut-limitation error from the data set. It is thus preferable to use in-gamut colors for the known *Lab* values at the start of the test. In-gamut *Lab* values can be obtained from measurement of a printed target. If we use only *Lab* values that have been obtained from measurement of a printed sample then by definition these *Lab* values are in-gamut. If this condition is met, then the ΔE calculation between known and measured *Lab* is due solely to errors in the profile and color conversion. In summary, the ΔE calculation shows the difference between a particular *Lab* color you wanted to reproduce and the *Lab* that you would get if you used that profile and that printer system.

To evaluate the A2B1 part of an output profile a *CMYK* target can be converted to *Lab* using the profile under test. In this test it is possible to reuse the *CMYK*-*Lab* from profile generation, or if there is concern about incestuous use of training versus testing data, new *CMYK*-*Lab* pairs for testing can easily be produced by printing and measuring any target. The procedure adopted here used the following process. Initially, to create a profile, some *CMYK* chart values are printed. When the *CMYK* patches are printed and measured they produce some (measured) *Lab*. The *CMYK*-*Lab* pairs are what is required for profile generation. If we use the A2B1 part of an output profile to convert *CMYK* values to (predicted) *Lab*, then the ΔE between measured and predicted *Lab* provides an indication of the error in the A2B1 part of the output profile.

Note that is possible to do a software only "round-trip" test. Round tripping involves taking some (optionally in-gamut) *Lab* values and converting them to *CMYK* and then back again to *Lab*. The ΔE between the start and finish *Lab* gives us an indication of the accuracy of the reversibility of a profile lookup table. The advantage of this test is that it can use separate data for training and testing and also it can be conducted entirely in software. The disadvantage of this test is that it does not separately measure the accuracy of the forward and reverse parts of the output profile, and it only tells us about the reversibility of the profile transform and not about the accuracy of the underlying device characterization. The round-trip test does not tell us much about the likely result when processing images—as a very poor profile can have excellent reversibility.

In the round-trip test, the starting *Lab* can be generated by converting *CMYK* values to *Lab* via the output profile under test. To obtain in-gamut values for this test, the same profile can be used to create the initial data set by processing *CMYK* values to *Lab* using the A2B lookup table, prior to conducting the round trip test. However, note that in this instance the profile is being used to create the test data that are going to be used to test the profile. Such incestuous processes should be avoided. A further issue with generating test data using this mechanism of "forced clipping" is that, by definition, considerable data are mapped to the edge of gamut region. Data values that are clustered in this edge of boundary region can skew the accuracy prediction.

PRINTER PROFILE—RESULTS

An output profile was made for an Epson Stylus Pro 4000 ink jet printer in *CMYK* mode using a ColorBurst 3.8 RIP with Epson Ultrachrome inks and Felix Schoeller proofing roll paper. In all cases the output profile was made from the same single measurement of the ECI 2002 target measured on a SpectroScan/Spectrolino. Default values were used in each program for all settings of black generation and profile quality/lookup table size. When printing to the Epson 4000 via the ColorBurst RIP, no ink limiting or linearization was used. The ColorBurst RIP is used merely to print to the device. In ICC terminology we may say that each vendor is asked simply to make a profile between the *CMYK* values sent to the printer and the *Lab* values that result from measurement of that target. Ink limiting and linearization is commonly used with ink jet printers to make better use of the available test patches. In this test the RIP was used merely to provide a mechanism to drive the device as a *CMYK* printer. The RIP was a fixed parameter and was not a variable in this experiment. In the printer profile testing we were interested in the accuracy of some device dependent *CMYK* values and their pairing with some CIE *Lab* values. As long as the RIP was repeatable that was more important than any RIP behavior, per se. In this testing we did not want to interfere with the data that are used in profile building. We feel it is unfair to present to profile makers data that have been preprocessed by a third party. The more parties involved in data interference the more potential for cumulative errors and for finger pointing and blame. We also strive to create a testing protocol that leaves no doubt as to what is being tested, so we do not have questions such as is the error in the RIP or the profile? We also believe that ink limiting and linearization may be a trade off between better behaved and smaller gamut data versus badly behaved and larger gamut. If we give data produced under linearized conditions to all vendors, we would expect different results, and probably all vendors would be expected to do better than published. The point is that as all vendors were given the same data, the results show the position of each vendor relative to each other, thus the published numbers are a valid comparison. Further, the rationale is that we do not want to necessarily make it easy for software, the tests seek to distinguish those products that are able to deal with raw printer data. Sophisticated software should be able to internally preprocesses that data so that vendors have enough information in the ECI 2002 target patches to internally analyze the data and linearize if necessary.

In this test we separately measure the accuracy of the A2B1 and B2A1 parts of the output profile and provide these results and also an average. To evaluate the B2A1 part of an output profile the *Lab* values of the measured ECI 2002 chart were placed in an image. We used a program written in our laboratory; however, a free program from GretagMacbeth called Logo ColorLab can be used for this purpose. Using Photoshop, the *Lab* image was converted to *CMYK* using each profile in turn. The ACE CMM was used and the intent selected in Image > Mode > Convert to Profile was Absolute Colorimetric. The *CMYK* image was printed and the *Lab* of each patch was measured. The measured *Lab* was compared to the *Lab* that was in the image being sent to the printer. The mean ΔE was calculated and averaged over all the patches. Because of the way the test was conducted all colors sent to the printer were in gamut. The test shows the difference between the particular *Lab* color that was intended for reproduction and the *Lab* that you would get if you used that printer profile and that printer. The error between the *Lab* values you wanted and the *Lab* values you achieved is calculated and shown in terms of ΔE.

To evaluate the A2B1 part of the output profile the ECI 2002 *CMYK* image was converted to *Lab* using each profile in turn. We know what *Lab* we had from the measurement file—when the ECI 2002 target was first printed and measured for profile generation, so if we use the A2B1 table of each output profile to predict the *Lab*, then the ΔE between the predicted *Lab* and the measurement file *Lab* tells us the error in the A2B1 part of the output profile. To do this test, the ECI 2002 *CMYK* chart image was opened in Photoshop and Image > Mode > Assign Profile was used for each profile in turn. The image was converted to *Lab* using Image > Mode > Convert to Profile (*Lab* Color). The ACE CMM was used and the intent was Absolute Colorimetric. Dither was not selected. The *Lab* of each patch in the digital file was averaged using a special program we have written. The ECI 2002 *CMYK* image is synthetically generated and thus contains no noise, unlike data used in the scanner analysis section. Finally, the values were compared to the measurement file. ΔE was calculated using GretagMacbeth MeasureTool. To eliminate ink jet print instability, all prints were allowed to stabilize for at least 24 h before measurement.

Keep in mind that the B2A1 table is used for processing/printing of images and is the more important column in Table VI. The A2B1 column is expected to be better than the B2A1 column as the A2B1 calculation involves only a software process, while the B2A1 involves printing and measuring that leads to greater inaccuracies.

The results produced by vendors with an average ΔE of around 2 are very good and are likely to produce excellent results in all printer based workflows. However, keep in mind that a large maximum ΔE has the potential to cause problems in particular image colors. The numbers reported here are in accordance with data presented by other researchers.[34] The number of grid points used in the profile can also be examined to confirm for example that the forward lookup table has more nodes while the reverse table tends to be more sparsely populated. In looking at the B2A1 table size we see that MonacoProfiler is disadvantaged in this comparison as their profile had a lower number of cube nodes (their default setting) in the lookup table compared to all other entries in the list. The Chromix ColorValet profile was made using a special target supplied by the vendor. The target was downloaded from the vendor's web site, printed on the Epson 4000 printer and sent by overright courier to Chromix. The ColorValet profile was created with a vendor-specified target but was subjected to testing based on the ECI 2002 target used in the main series of our tests.

In the lower part of the results table an entry is shown for the ColorBurst RIP 3.8 Generic Profile Epson Premium Semi-Matte Paper. To generate this entry, the printer environment (ink limits and linearization) were set in accordance with how the generic profile was made, but the paper used was different. The profile assumed Epson Premium Semi-Matte paper while we used Felix Schoeller Semi-Matte paper. This entry is shown for the situation where a user may have a printer and paper but no profile. The user needs a generic profile to complete their workflow. Instead of custom profiling, an option open to the user is to use the closest available generic profile. The entry shown in the table is the

Table VI. Results for the quality of the B2A and A2B parts of the printer profile.

Printer profile quality Epson Stylus Pro 4000 with ColorBurst RIP 3.8	B2A1		A2B1		Average ΔE
	Mean/Max ΔE	Lookup table size	Mean/Max ΔE	Lookup table size	
Heidelberg PrintOpen 5.1 (Windows)	2.32 (10.82)	33	1.11 (8.42)	16	1.72
Fujifilm ColourKit Profiler Suite 4.2	2.56 (14.83)	33	1.07 (9.78)	17	1.82
GretagMacbeth ProfileMaker 5.0.1	3.26 (12.40)	25	1.12 (6.27)	11	2.19
ColorSolutions basICColor Print4c 2.1	3.02 (14.30)	33	1.49 (11.39)	17	2.26
TGLC PerfX Color Management 1.2.8	3.64 (14.00)	33	1.69 (8.39)	17	2.67
X-Rite Monaco Profiler 4.7	4.17 (19.72)	17	1.54 (9.85)	17	2.86
CHROMIX ColorValet Print 2.3	3.75 (12.84)	25	3.22 (20.34)	11	3.49
ColorBurst RIP 3.8 Generic Profile Epson Premium Semi Matt Paper[a]	4.05 (10.73)	21	Not applicable	17	Not applicable
GMG ColorProof 04	4.57 (19.51)	Not applicable	Not applicable	Not applicable	Not applicable

[a]Generic profile supplied by ColorBurst/Epson but intended for Epson paper, we used Felix Schoeller paper, which is similar but not the same.

accuracy the user can expect in this situation. The result for GMG ColorProof result was obtained external to this testing. The GMG system is a personal computer based system and directly controls the Epson 4000 and the SpectroScan. The measurements of the ECI 2002 target used in the main series of tests formed the "aim" or "target" values for the GMG system. GMG ColorProof took control of the Epson printer and iteratively created an ECI 2002 target that was measured and compared to the target values. The data shown here for GMG are not typical of this system, the data are much worse than normally expected.

PRINTER PROFILE—VENDOR'S PREDICTIONS

We were pleased to see that two vendors provide feedback following profile generation. When we see a big disparity between our calculations and the vendor predictions we ask the question "are we conducting the same test?." Further we ask the question "are vendors doing a realistic, meaningful test?." In other words, "Are vendors using a populated, hence, (quantized) lookup table to predict the values?." The conclusion from Table VII is that there is little correlation between what vendors are currently reporting for the accuracy of an output profile compared to values obtained in this study via empirical methods.

CONCLUSIONS

A simple, easy to compute metric for input profile quality is described and evaluated for a number of commercial profiling software applications. The ΔE for profiles made in this research was in the range of $0.60 \leqslant \Delta E \leqslant 2.46$. It is reassuring to note that the vendor's predictions agree with those calculated in this research, which suggests that researchers and vendors are converging to a universally agreeable test protocol. Analysis presented here shows that some vendors are using the same core technology in both professional and consumer products. Historical analysis of input profile accuracy measurements shows that commercial profiling products have been considerably improved in recent years. While the described colorimetric methodology does not explicitly test real images, examination of the ΔE error values for an input profile does reveal information about the potential behavior of the profile and provides a simple straightforward measure that is comprehensible to all levels of the color management user community.

A test for monitor profiles was described that included evaluating the gamma, white point, and accuracy of color reproduction for 24 chosen color patches. The commercial products were able to reproduce 24 patches of a color checker with an average ΔE of 2.92–5.81. The results show a reduced interest by profile makers to create traditional gamma values that were more suitable to CRT devices.

There is a great deal of interest in evaluating printer profiles for direct printing of images, there is also interest in using ink jet printers in color management proofing workflows. The average accuracy of the output profile colorimetric intent was shown to be have a range of $1.72 \leqslant \Delta E \leqslant 3.49$. In a printer profile there are three possible metrics:

Table VII. Results for the quality of the printer profile compared to vendor's prediction.

Printer profile quality Epson Stylus Pro 4000 ColorBurst RIP 3.8		Mean (Max) ΔE A2B1 used for proofing
X-Rite Monaco	Our findings	1.54 (9.85)
Profiler 4.7	Vendor data	0.26 (1.3)
Fujifilm ColourKit	Our findings	1.07 (9.78)
Profiler Suite 4.2	Vendor data	0.15 (6.47)

accuracy of the B2A1 tag (PCS to Device); A2B1 tag (Device to PCS); and a round trip test. It was shown that the round trip test is less useful.

The data presented here are fundamentally a description of methods to estimate the colorimetric accuracy of profiles. The use of commercial profiling products is only to illustrate how the methodology may be used. An important theme in this research is a modular approach that separately evaluated the quality of scanner, monitor, and printer profiles.

The appearance of images is an important criterion that must form an integral part of any profile evaluation process. There are also other attributes such as smoothness of the transform that should be considered in conjunction with the colorimetric testing described here. Researchers are working in this area and future metrics are expected to incorporate other aspects of visual image quality.

Color management is an important area as the increasing number of digital workflows fuels the demand for accurate, reproducible color in an open-loop color management system. This research is useful in helping the graphic arts industry achieve better quality profiles through a system of standards, which, hopefully, will lead to a greater acceptance of ICC-based workflows. It should be noted that the issue of profile quality is being addressed by a newly formed ICC group—Profile Assessment Working Group.

ACKNOWLEDGMENTS

The author is grateful to the vendors who have collaborated and allowed inclusion of their products in this review and who provided valuable feedback and suggestions for the test procedures. It is a pleasure to acknowledge many discussions with my knowledgeable colleague Dr. Paul D. Fleming, Western Michigan University. The author is extremely grateful to Steve Rankin, Doug Baker, and Dwight Lemmon at X-Rite for providing access to instrumentation for the monitor profile testing part of this work. The research has been partially funded by National Science Foundation Major Research Instrumentation Grant No. NSF MRI 0215356.

REFERENCES

[1] *Recent Progress in Color Management and Communications*, edited by R. Buckley (IS&T, Springfield, VA, 1998).

[2] E. J. Giorgianni and T. E. Madden, *Digital Color Management* (Addison-Wesley, Reading, MA, 1998).

[3] T. Johnson, "An effective colour management architecture for graphic arts", Proc. TAGA, 88 (2000).

[4] A. Sharma, *Understanding Color Management* (Delmar Thomson, New York, 2004).

[5] International Color Consortium, Specification ICC.1:1998-09, File format for color profiles, www.color.org. (Specification ICC.1:2001-12, Version 4.0.0).

[6] R. A. Holub, "End-to-end analysis of color errors in a soft-proofing system", Proc. TAGA, 177 (2000).

[7] A. Sharma, M. P. Gouch, and D. N. Rughani, "Generation of an ICC profile from a proprietary style file", J. Imaging Sci. Technol. **46**, 26 (2002).

[8] R. S. Berns, "Science of digitizing paintings for color-accurate image archives: A review", J. Imaging Sci. Technol. **45**, 305 (2001).

[9] J. Orava, T. Jaaskelainen, and J. Parkkinen, "Color errors of digital cameras", Color Res. Appl. **29**, 217 (2004).

[10] H. Zeng and M. Nielsen, "Color transform accuracy and efficiency in ICC color management", *Proc. IS&T/SID 9th Color Imaging Conference* (IS&T, Springfield, VA, 2001) p. 224.

[11] H. Zeng, "Color accuracy in ICC color management system", *Proc. IS&T PICS* 2002 (IS&T, Springfield, VA, 2001) p. 175.

[12] R. M. Adams and J. B. Weisberg, *GATF Practical Guide to Color Management* (GATF Press, Pittsburgh, PA, 2000).

[13] W. Kress, "Digitization and metric conversion for image quality test targets", *Proc. IS&T PICS 2003* (IS&T, Springfield, VA, 2003) p. 82.

[14] R. S. Berns and D. M. Reiman, "Color managing the third edition of Billmeyer and Saltzman's Principles of Color Technology", Color Res. Appl. **27**(5), 360 (2002).

[15] H. Boll, Matlab Script available at www.mathworks.com/matlabcentral/fileexchange/loadFile.do?objectId=3674&objectType=FILE.

[16] P. Zolliker, M. Daetwyler, and K. Simon, "On the continuity of gamut mapping algorithms," Proc. SPIE **5667**, 220 (2005).

[17] H. Büring, P. G. Herzog, and E. Jung, "Evaluation of current color management tools: Image quality assessments", *Proc. Second CGIV* (IS&T, Springfield, VA, 2004).

[18] C. Kuo and Y. Ng, "Perceptual color contouring detection and quality evaluation using scanners", J. Imaging Sci. Technol. Vol. **49**, 41 (2005).

[19] F. Sigg, *A New Tool for Quantitative Comparison of Color Differences, Test Targets 4.0* (School of Print Media Publication, RIT Rochester, NY, 2004).

[20] R. Fisch and S. Bartels, "A colorimetric test for reflection CMYK colorant output", Proc. TAGA, 204 (1999).

[21] D. Rich, *Graphic Technology—Improving the Inter-Instrument Agreement of Spectrocolorimeters* (Committee for Graphic Arts Technologies Standards (CGATS) NPES, Reston, VA, 2004).

[22] H. R. Kang, *Color Technology for Electronic Imaging Devices* (Bellingham, WA, 1997), p. 55.

[23] G. Sharma and H. J. Trussell, "Digital color imaging", IEEE Trans. Image Process. **6**, 901 (1997).

[24] H. R. Kang, "Color scanner calibration", J. Imaging Sci. Technol. **36**, 162 (1992).

[25] X. Rong, P. D. Fleming, and A. Sharma, "Quantitative analysis of ICC profile quality for scanners", Proc. TAGA, 3 (2004).

[26] A. Sharma and P. D. Fleming, "Evaluating the quality of commercial ICC color management Software", Proc. TAGA, 336 (2002).

[27] A. Sharma and P. D. Fleming, "Measuring the quality of ICC profiles and color management Software", Seybold Report **2**(19), 3 (2003).

[28] A. Sharma, "Measuring the quality of ICC profiles and color management software", Seybold Report **4**(20), 10 (2005).

[29] W. Kress, "Digitization and metric conversion for image quality test targets—Part II", Proc. SPIE, **294**, 39 (2004).

[30] R. S. Berns, "Methods for characterizing CRT displays", Displays **16**, 173 (1996).

[31] Apple Computer, *What's New in ColorSync 2.5* (Apple Technical Publications, Cupertino, CA, 1998).

[32] E. A. Day, L. Taplin, and R. S. Berns, "Colorimetric characterization of a computer-controlled liquid crystal display", Color Res. Appl. **29**(5), 365 (2004).

[33] T. Sugiyama and Y. Kudo, "The display profiling method with eyes and colormatching accuracy in soft proofing", Proc. SPIE **5667**, 385 (2005).

[34] K. Kimura, T. Nakagawa, F. Nakasai, H. Takahshi, N. Utsumi, and S. Kitaoka, "Improvement of color matching accuracy with an ink jet proofer", *Proc. IS&T/SID 11th Color Imaging Conference* (IS&T, Springfield, VA, 2003) p. 238.